This manuscript has been accepted for publication in the Quarterly Journal of Experimental Psychology.

The role of social attention in older adults' ability to interpret naturalistic social scenes

Sarah A. Grainger[1], Henriette R. Steinvik[1], Julie D. Henry[1] & Louise H. Phillips[2]

[1]School of Psychology, University of Queensland

[2]School of Psychology, University of Aberdeen

Correspondence should be addressed to: Dr Sarah Grainger, email: s.grainger@uq.edu.au, Address: School of Psychology, University of Queensland, St Lucia, Queensland, 4072, Australia.

Abstract

Age-related differences on theory of mind (ToM) tasks are well established. However, the literature has been criticised for predominantly relying on tasks with poor ecological validity, and consequently it remains unclear whether these age differences extend to tasks with greater realism. In addition, we currently have a limited understanding of the factors that may contribute to age-related declines in ToM. To address these issues, we conducted two studies that assessed age differences in ToM using multimodal social scene stimuli. Study 1 also examined eye-movements to assess whether biases in visual attention may be related to age-related difficulties in ToM, and Study 2 included an assessment of social attention (as indexed by biological motion perception) and working memory to assess whether these capacities may explain age difficulties in ToM. In both studies, the results showed that older adults performed worse than their younger counterparts on the ToM tasks, indicating that age-related difficulties in ToM extend to measures that more closely represent everyday social interactions. The eye-tracking data in Study 1 showed that older adults gazed less at the faces of protagonists in the social scenes compared to younger adults, however these visual biases were not associated with ToM ability. Study 2 showed that older age was associated with a reduced ability to detect biological motion cues, and this mediated age-related variance in ToM ability. These findings are discussed in relation to competing theoretical frameworks of ageing that predict either improvements or declines in ToM with age.

Keywords: ToM, Ecological Validity, Eye-tracking, Biological Motion Perception, Working Memory

Social cognition refers to the ability to decode, understand, and react to the social cues sent out by other people, and is a critical predictor of social competency, well-being and mental health at all stages of the human lifespan (Grühn, Rebucal, Diehl, Lumley, & Labouvie-Vief, 2008; Henry, von Hippel, Molenberghs, Lee & Sachdev, 2016; Phillips, Scott, Henry, Mowat & Bell, 2010). It is therefore of concern that research involving older adults has consistently found evidence for age-related decline in core social cognitive skills. One component of social cognition that consistently shows age-related decline is *theory of mind* (ToM), which refers to the ability to attribute intentions and more complex emotional and mental states to others (for a meta-analysis see Henry, Phillips, Ruffman, & Bailey, 2013; Moran, 2013).

However, a limitation of the ToM literature is that, with few exceptions, the measures that have been used to identify age effects have had only limited ecological validity. For instance, most ToM measures require participants to infer mental states from brief stories or images of isolated eye regions (Henry et al., 2013). While these paradigms are useful because of their high level of experimental control, they are limited by the fact that they are not an accurate representation of how social interactions are experienced in real life (i.e., they only present social cues via a single modality). Therefore, it is possible that the use of these paradigms may overstate the difficulties that older adults may experience in their everyday lives. Not surprisingly, ecological validity has been identified as a critical consideration in the social cognition literature to date and researchers have emphasised the need for more ecologically valid research designs (Isaacowitz & Stanley, 2011; Kunzmann & Isaacowitz, 2017).

According to Brunswik (1955), ecological validity refers to the degree to which the characteristics of an experimental design represent the phenomenon of interest as it occurs in the real world, and it can be evaluated in relation to several aspects of the research process (e.g., the research setting, the technique used to elicit a behavior, the stimuli). Because peoples' social worlds differ based on factors such as culture, social circles, and hobbies, designing tasks that are completely representative for participant samples that include extreme groups (i.e., older and younger) is a particular challenge. Thus, the goal is to attempt to incorporate tasks that provide a reasonably close representation of social interactions that both younger and older adults may experience in their day-to-day social lives.

One well-validated measure that assesses ToM ability using naturalistic social scene stimuli designed to resemble day-to-day social interactions is The Awareness of Social Inference Test (TASIT, McDonald, Flanagan & Rollins, 2011). The TASIT stimuli include dynamic everyday social interactions between multiple protagonists, making it a more realistic and thus more ecologically valid assessment of ToM than traditional measures (Achim, Guitton, Jackson, Boutin, & Monetta, 2013). Participants are required to identify the mental state of a key protagonist, assessing their intention when they made a particular statement. Although this measure was originally designed to measure social perceptual deficits in clinical groups, two previous studies have administered this measure to older adult populations and found that it is sensitive to age-related decline (Burdon, Dipper & Cocks, 2016; Phillips et al., 2015). However, these two studies used UK samples, and the protagonists in the TASIT stimuli are Australian. Burdon et al. highlight that it is important to establish whether age effects on this task also occur in an Australian sample. Indeed, understanding social interactions portrayed by people with an unfamiliar accent may impose demands that potentially reduce older adults' ability to do well on the task. Therefore, in these studies we look at age differences in the TASIT in both Australian (Study 1) and UK (Study 2) samples to see whether the age pattern replicates.

Lifespan motivational models, such as the selective engagement hypothesis (Hess, 2006) suggest that older adults become more selective in how they invest their resources. Thus, selection represents an adaptive motivational process in response to age-related changes, including capabilities and developmental goals (for reviews, see: Carstensen, Mikels, & Mather, 2006; Hess, 2006). Consistent with this perspective, evidence suggests that, relative to their younger counterparts, older adults are more likely to exert cognitive effort in situations of high relevance and meaningfulness (e.g., Germain & Hess, 2007; Hess, Germain, Rosenberg, Leclerc, & Hodges, 2005; Hess, Germain, Swaim, & Osowski, 2009; Hess, Rosenberg, & Waters, 2001). Therefore, it might be expected that compared to the stimuli used in *traditional* ToM tasks, older adults may be more motivated to engage with stimuli that they can relate more to (i.e., stimuli that resemble everyday social interactions), and as such may perform better with the provision of these stimuli.

However, there may be competing negative influences on older adults' ability to process more ecologically valid social cues. Many cognitive ageing models (Craik & Byrd, 1982; Glisky, 2007; Salthouse, 1995; Ruffman et al., 2008) highlight the neural and cognitive changes that inevitably occur with normal adult ageing. Although older adults might be more motivated to engage with stimuli that are more representative of real life, they might also experience greater difficulties processing these stimuli because of the increased perceptual complexity in realistic stimuli. That is, the increased socio-emotional and background cues afforded by dynamic social stimuli may be more demanding on older adults' limited neural and cognitive resources, and this may in fact disrupt their performance on tasks with higher ecological validity.

In the current studies, we will examine three possible ways in which cognitive and attentional processes might influence age differences in processing naturalistic social cognitive cues. The first is whether older adults' tendency to bias attention away from the most socially salient information in a scene (i.e., faces) could be important (Study 1): this issue is outlined in detail below. The second is to explore possible links between general working memory capacity and age effects on processing naturalistic social cues (Study 2).

The third possibility is that the ability to attend to and rapidly process basic social cues is likely to be important in age differences in attributing mental states. The most common method of assessing attentional capacity to process basic social cues is through biological motion tasks, which often use point-light displays of body parts, and ask for decisions about human agency. Older adults are less able to rapidly process biological motion cues (e.g. Billino, Bremmer & Gegenfurtner, 2008), and these age differences in biological motion processing may be important in understanding age effects in broader social cognition (for more detail, see Study 2).

As mentioned earlier, one particular aspect of older adults' changing cognition that may be important for social understanding, is the tendency to focus visual attention on particular aspects of a complex array. Visual biases to social cues may be a potentially important consideration when interpreting age effects in social cognitive function. Research in the emotion perception literature has consistently shown that older compared to younger adults show a bias in the way they visually attend to emotional faces. Specifically, older adults attend less to the eye region of an emotional face compared to younger adults (He, Ebner, & Johnson, 2011; Murphy & Isaacowitz, 2010; Sullivan, Ruffman, & Hutton, 2007; Wong, Cronin-Golomb, & Neargarder, 2005), and this bias away from the most emotionally rich area of the face has been argued to be a contributing factor to older adults' difficulties in decoding emotional expressions (Wong et al., 2005). However, the relationship between visual attention biases and emotion perception accuracy remains unclear. While two studies have shown that older adults' reduced attention to eyes is related to their ability to decode an emotion (Wong et al., 2005; Sullivan, Campbell, Hutton & Ruffman, 2015), four others have failed to find any such relationship (e.g., Circelli, Clark & Cronin-Golomb, 2013; Grainger, Henry, Phillips, Vanman & Allen, 2017; Sullivan et al., 2007). Moreover, because nearly all of the studies in this literature have used static photos of emotion, they have limited ecological validity, and therefore the visual biases may not be representative of how older adults actually attend to social information in everyday real life.

Another limitation of this literature is the exclusive focus on emotion perception, and not other social cognitive processes such as ToM decisions. To date, only one ageing study has provided an assessment of eye gaze patterns to dynamic social interactions (Vicaria, Bernieri & Isaacowitz, 2015). In this study, participants watched videos depicting two people discussing a controversial topic, and then had to judge the rapport between the two protagonists while their eye movements were being monitored. The results indicated that older and younger adults differed in the way they attended to the interactions, such that older adults focused less on the faces of the protagonists relative to younger adults. This finding suggests that older adults' attentional bias when looking at photographs of faces, which is well established in the emotion recognition literature, may extend to dynamic social interactions. Interestingly, Vicaria et al. found that gazing to the faces of the protagonists was associated with greater ability to judge rapport in young adults but poorer ability to judge rapport in older adults, suggesting that young and older adults may benefit from different visual approaches when viewing more complex social information.

However, while Vicaria et al. (2015) provides an important first step in understanding how older adults attend to dynamic social interactions, only one specific type of social interaction was assessed, and all background information was kept constant across interactions. As a consequence, the background contextual information was not useful in facilitating social judgments. Given that social interactions often depend heavily on the context in which they are embedded, it is important to understand if age-related biases in visual attention occur during ToM processing with the provision of realistic context rich scenarios. Indeed, use of such scenarios appears to be particularly critical in any assessment of potential age differences because, relative to their younger counterparts, older adults have been shown to place a greater reliance on contextual cues than internal representation to support their information processing (Lindenberger & Mayr, 2014; Smith, Park, Earles, Shaw, & Whiting, 1998; Spieler, Mayr, & Lagrone, 2006). In many cognitive domains – such as memory, perception and action management – age-related deficits have been found to decrease or disappear when task-relevant contextual cues are provided (Lindenberger & Mayr, 2014; Smith et al., 1998). Moreover, *irrelevant* context (e.g., distracting cues that are unrelated to the goal of a task) have been found to disproportionately disrupt older adults' performance, presumably because of older adults' greater prioritisation of contextual information (Gutchess et al., 2007; Lindenberger & Mayr, 2014).

Indeed, two studies in the emotion perception literature have shown that contextual objects and scenes surrounding a target emotional face can influence older adults' emotion perception abilities (Ngo & Isaacowitz, 2015; Noh & Isaacowitz, 2013). In both of these studies, participants were asked to identify emotional facial expressions presented with a congruent background (i.e., same emotional context and target emotion) or an incongruent background (i.e., different emotional context and target emotion). The results showed that for the incongruent trials, older adults made more errors than younger adults, whereas for the majority of congruent trials, age differences in accuracy were eliminated. Ngo and Isaacowitz (2015) suggested that older adults might show a greater tendency to rely on spontaneous encoding of context, while younger adults are more likely to engage their top-down attentional control to ignore irrelevant context. Consistent with this possibility, Noh and Isaacowitz (2013) showed using eye tracking that older adults attended more to contextual cues (i.e., emotional laden objects and body posture context) relative to younger adults. However, while these studies have demonstrated how contextual cues influence emotion processing, it remains unclear whether higher order social cognitive processes such as ToM can also benefit from increased contextual information. In addition, another limitation of both of these studies was the use of exclusively static stimuli. Therefore, an important next step in this literature will be understanding how the presentation of additional contextual information during more naturalistic stimuli influences higher order social cognitive processes such as ToM.

As previously mentioned, the TASIT is a well-validated social cognition measure that provides an assessment of ToM ability using more naturalistic social scene stimuli that are representative of everyday social encounters. Another strength of this measure is that it includes two separate ToM tasks that include stimuli with either minimal or enriched contextual information, allowing the influence of contextual information on ToM processing to be directly compared. It is important to note that the previous studies that examined ageing and TASIT performance either only assessed one component of the TASIT (Phillips et al., 2015) or did not make a direct comparison as to whether the age effects were influenced by context (Burdon et al., 2016). Therefore, further research is required to directly compare the influence of contextual cues on ToM processing.

Study 1

Study Aims

Study 1 had two primary aims. The first was to examine ToM capacity in older and younger adults using naturalistic stimuli (TASIT videos) that more closely represent every-day social encounters. These assessments will allow for the replication of the previously reported age-related findings on the TASIT using an Australian sample (Burdon et al., 2016; Phillips et al., 2015), but most critically, these assessments will examine the role of contextual processing specifically. The second aim was to provide the first assessment of visual biases to complex mental state social scenes to establish whether older adults show biases in visual attention when presented with social stimuli that are more representative of everyday life. Most importantly, the assessment of eye-gaze patterns will allow us to establish whether visual biases are related to older adults' performance on the ToM assessments.

With respect to specific predictions for ToM performance, based on prior findings (Burdon et al. 2016; Phillips et al., 2015), we predicted that older adults would perform more poorly than younger adults on both measures of ToM understanding (TASIT 2 and TASIT 3). However, of greatest interest was whether age differences would be equivalent across these two measures. This is because, whereas TASIT 2 incorporates only paralinguistic cues, TASIT 3 additionally incorporates contextual cues (for instance, a visual cue that indicates the true state of affairs, or a prologue that reveals the protagonist's true thoughts). Consequently, if older adults do place greater reliance on contextual information to facilitate ToM understanding, any age differences on TASIT 3 should be smaller relative to TASIT 2. Conversely, if these additional contextual cues simply increase task complexity, age-related difficulties should be greater on TASIT 3.

With respect to eye-gaze patterns, based on prior findings (Vicaria et al. 2015), we expected older adults to gaze less at the faces of protagonists relative to younger adults. However, the predictions regarding the relationship between eye-gaze patterns and ToM ability are less clear. On the one hand, it could be expected that gazing to the more socially rich areas of the scenes (i.e., faces) will be associated with greater ToM performance in both young and older adults. However, on the other hand, it is possible that increased gazing to the faces may be not be beneficial for older adults. This is because the only other study to date that has examined older adults' eye-gaze patterns in response to dynamic stimuli (Vicaria et al., 2015) found that increased gazing to the faces of protagonists was associated with poorer judgment accuracy in older adults.

Method

Participants

Fifty older adults (M_{age} = 75.38, *SD* = 6.23, age range: 65-94 years, 33 female) and 48 younger adults (M_{age} = 20.67, *SD* = 8.82, age range: 17-28 years, 33 female) participated in this study[1]. Older adults were recruited via local advertisements and from a research register at the University of Queensland, and were reimbursed $40 for their participation. Younger adults were recruited from the first-year psychology participant pool or via a paid participant pool at the University of Queensland, and received either course credits or $40, respectively. To be eligible, all participants had English as their first language, 20/30 or corrected vision, and no history of substance abuse or presence of a psychiatric or neurological disorder. All older adults were screened for abnormal cognitive decline using the Australian version of Addenbrooke's Revised Cognitive Examination (ACE-R; Mioshi, Dawson, Mitchell, Arnold, & Hodges, 2006), and scored above the recommended cut-off of 83 out of a possible 100.

As can be seen in Table 1, there were no differences between the two groups with respect to years of education, or self-rated health. Older adults reported lower negative affect as indexed by scores on the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), and higher full-scale IQ as indexed by performance on the National Adult Reading Test (NART; Nelson, 1982).

Measures

Dynamic ToM tasks. Parts 2 and 3 of the TASIT (Form A; McDonald et al., 2011) were used to assess ToM ability. These include audio-visual videotaped vignettes (15-60 seconds each) that depict everyday social interactions enacted by unfamiliar Australian actors, and examine the ability to make inferences about others' intentions, attitudes, and meanings. Importantly, these stimuli incorporate key paralinguistic cues (i.e., voice tone,

---

[1] 15 older (10 female) and 14 younger (8 female) adults in this sample were administered a placebo (saline) nasal spray prior to completing the ToM assessments as part of a separate larger study.

gestures, body posture) that are essential in order to differentiate between social inferences (i.e., when someone is being sincere vs sarcastic). The protagonists in these videos are young and middle-aged, males and females, who are involved in a certain type of social relationship (i.e., a couple, two friends, work colleagues).

Part 2: Social Inference – Minimal, includes 15 vignettes that depict an interaction between two protagonists. Five vignettes include sincere remarks, in which the meaning of what the target speaker says is consistent with paralinguistic cues, e.g., ¨*Michael: Sorry, I can't take that class I said I'd take on Friday. Ruth: (sincerely) That's OK, I know you're busy. Don't worry about it*¨. Another five include simple sarcasm, in which the actors' paralinguistic cues reveal the sarcasm that is meant to be understood non-literally, e.g., ¨*You have been a great help!*¨. The remaining five vignettes include paradoxical sarcasm, in which the interaction makes sense only if it was understood that one person was being sarcastic, e.g., ¨*Gary: Are you sure you've got your passport? Keith: (sarcastically) Oh, yes, I tore it up and threw it away. Gary: Good, that's okay then*¨. Comprehension is assessed through four questions for each vignette, covering four facets of understanding about the protagonists, including thoughts (what s/he knows), feelings (what s/he feels), intentions (what s/he intends to do) and meaning (what s/he means by what is said). Response options include *yes*, *no*, and *don't know*. A total score (out of 60) was created, and then converted into a percentage score.

Part 3: Social Inference – Enriched, includes 16 vignettes, in which visual (i.e., background cue) and verbal (i.e., prologue or epilogue) contextual cues are provided about the characters' true knowledge and beliefs, in addition to paralinguistic cues as in TASIT 2. The visual contextual cues are provided by a camera shot of a cue that reveals the true state of affairs (e.g., when a man is asking his wife if their son had finished his dinner, the camera is focused on the son's dinner plate to show whether or not he had finished his dinner). The verbal contextual cues are typically included as a prologue or epilogue that reveals how someone is truly feeling (e.g., In the first scene, a woman is chatting to her friend about how she believes her husband has gained weight, and in the following scene she tells her husband that she doesn't think he has gained weight). Thus, an important point is that for TASIT 3, the contextual cues reveal the true state of affairs, therefore attending to and understanding these cues is essential to understand the mental states of the protagonists in the scenes. Half of the vignettes incorporate scenes that involve literally untrue remarks (i.e., lies), e.g., *¨Of course you don't look fat¨* (visual cue shows that he actually is looking fat). The remaining vignettes consist of sarcastic remarks (e.g., saying that he doesn't look fat but using a sarcastic manner). Similar to Part 2, a total score (out of 64) was created, and then converted into a percentage score.

Apparatus

To create separate video files for each of the TASIT videos, *iMovie* (Version 10.0.9) video editing tool was used, and *Wondershare video converter* (Version 4.4.5) was used to convert these into mp4 files. The *SensoMotoric Instruments* (SMI Gmbh, 2015) *Experiment Centre*^{TM} program (Version 3.5) was used to present each of the TASIT components, which were displayed on a *Dell*^{TM} 56cm (22") Widescreen Flat Panel Monitor, with a screen resolution of 1680x1050 pixels.

To collect eye-gaze data, a SMI RED500 remote mobile eye tracker (positioned below the monitor) with *iViewX*^{TM} (Version 2.8, 2011) software was interfaced and controlled via a Dell Latitude Laptop. This system is non-invasive, with little indication that eye movements are recorded, as it allows free head movement. It also works with most glasses and contact lenses, and tracks both eyes, to an accuracy of 0.5 degrees, sampled at 500 Hz.

Remote eye-trackers provide a naturalistic method of measuring visual attention because they do not require the participant to keep their head constrained to a chin rest. However, one limitation of this technique is that participants are able to freely to move their head around and these movement artifacts can often result in the eye tracker being unable to capture eye movements during part of a trial. While we instructed participants to keep their head as still as possible during trials, some movements were not possible to prevent (e.g., sneezing or coughing) and therefore resulted in a small number of trials having missing eye gaze data. Participants with more than 30% missing eye gaze data for a particular trial were removed from the relevant analyses. In total, data from two younger adults and 13 older adults were eliminated from specific analyses for this reason or because of poor eye calibration, and this is comparable with previous eye-tracking studies (e.g., Isaacowitz, Wadlinger, Goren & Wilson, 2006; Murphy & Isaacowitz, 2010; Vicaria et al., 2015)

Procedure

The study received approval from The University of Queensland ethics committee. All participants were tested individually for approximately two hours, inclusive of a short break. Upon arrival, participants were given an information sheet, followed by a demographic questionnaire. The ACE-R was then administered to older adults. Next, all participants completed the NART.

To complete the ToM assessments, participants were led into an adjacent room, and seated in a comfortable armchair, with their eyes positioned centrally and approximately 60 cm from the monitor. Each participant then completed one of four counterbalanced orders of the tests that included assessments of emotion recognition ability that are not reported here. While viewing the stimuli, eye-movements were monitored using the remote eye tracker, and a nine-point calibration and validation of participants' eyes were conducted to ensure that the positional eye data were consistent. Following eye validation for each of the tasks, written instructions appeared on the screen, followed by a practice session, which included familiarization with the computer display and keyboard, as well as guidance through one practice trial.

Each experimental trial began with a fixation cross presented in the middle of the screen, whereby participants had to press the keyboard spacebar for the video to begin playing. After each video, the participant was asked four questions which appeared separately on the screen. When the participant was ready to respond, they made their response verbally to the experimenter who then recorded their responses in an Excel spreadsheet. There were no time limits to make responses.

Upon completion of the ToM assessments, participants were led back to the adjacent testing room to complete the Hospital Anxiety and Depression Scale (Zigmond & Snaith, 1983), which was presented online using the Qualtrics™ Software. Participants were then provided with reimbursement and debriefed.

To analyze the eye-tracking data, regions of interest (ROIs) were set up for each stimuli set using the SMI Gmbh (2015) *Begaze™* (Version 2.4) software. Following previously reported methods (Riby & Hancock, 2009; Vicaria et al., 2015), the face regions of the protagonists were marked with an oval shape while the body region was defined by the outline of the body excluding the face. After creating the ROIs, all videos were carefully checked to ensure that the ROIs adequately captured the face and body regions throughout the video. Percent fixation time was calculated within each ROI for each video scene, and used as the primary eye gaze measure. Percent fixation time reflects the percentage of fixations within an ROI relative to the total trial duration (Murphy & Isaacowitz, 2010). When a participants' gaze is not captured in any of the predefined ROI's, it means that their gaze is directed towards other areas of the computer screen that are not covered by the ROI's (e.g., any of the background scene). In the present analyses, a fixation was defined as an eye movement directed within a 1° visual angle for at least 100ms (Manor & Gordon, 2003).

Analyses

Mixed design analyses of variance (ANOVAs) were conducted for both the assessments of ToM ability and visual attention biases. For all of these analyses, *p* values were adjusted using Greenhouse-Geisser correction for violations of the sphericity assumption. Given that the focus of this study was on social cognitive ageing, significant main effects or interactions that did not involve age were not followed up formally. Finally, Pearson correlations were calculated to assess any potential relationships between visual biases and ToM performance. Data from one older adult in TASIT 2 is missing due to an experimenter error during data collection process.

Results and Discussion

Dynamic ToM Tasks

To assess the effect of additional contextual cues on mental state understanding, data were analyzed with a 2 x 2 mixed ANOVA with the between-subjects variable of *age group* (younger, older) and the within-subjects variable of *context* (Minimal: TASIT 2, Enriched: TASIT 3). There was a main effect of age group, $F(1, 95) = 75.86$, $p < .001$, $\eta_p^2 = .45$, denoting overall poorer performance for older ($M = 79.56$, $SE = .83$) relative to younger adults ($M = 89.87$, $SE = .84$). There was a main effect of context $F(1, 95) = 27.44$, $p < .001$, $\eta_p^2 = .22$, and an interaction emerged between age-group and context, $F(1, 95) = 4.73$, $p = .032$, $\eta_p^2 = .05$, which was followed up by examining the effect of context at both levels of age-group. This revealed a main effect of context for younger, $F(1, 47) = 4.85$, $p = .033$, $\eta_p^2 = .09$, and older adults, $F(1, 48) = 26.67$, $p < .001$, $\eta_p^2 = .36$. However, it can be seen that the effect size is much greater for older relative to younger adults, indicating that context had a greater impact on older adults' performance. These findings suggest that additional contextual information did not provide any benefit and possibly increased the difficulty of the task for older adults. Older and younger adults' scores are presented separately for each task in Figure 1.

Importantly, the observed age-related difficulties in both measures of ToM indicates that older adults have difficulties decoding mental states even when the stimuli are more naturalistic and representative of everyday social encounters. These findings are consistent with the two prior ageing studies from the UK that found age-related decline on the TASIT (Burdon et al., 2016; Phillips et al., 2015), and indicate that older adults have difficulties in ToM processing even when the stimuli are more familiar and representative of their real life social interactions.

Age-Related Visual Biases

To examine the eye-tracking data, we ran a 2 x 2 x 2 mixed ANOVA with the between-subjects variable of *age-group* (young, old) and the within-subjects variables of *context* (minimal context: TASIT 2, enriched context: TASIT 3), and *ROI* (faces, bodies). There were main effects of age-group, $F(1, 81) = 8.88$, $p = .004$, $\eta_p^2 = .10$, context, $F(1, 81) = 133.96$, $p < .001$, $\eta_p^2 = .62$, and ROI, $F(1, 81) = 314.47$, $p < .001$, $\eta_p^2 = .80$, and two way interactions between age-group and ROI, $F(1, 81) = 6.08$, $p = .016$, $\eta_p^2 = .07$, and between context and ROI, $F(1, 81) = 103.95$, $p < .001$, $\eta_p^2 = .56$. The three-way interaction between age-group, context and ROI was not significant, $F(1, 81) = 1.82$, $p = .181$, $\eta_p^2 = .02$.

The two-way interaction between age-group and ROI was followed up by testing the effect of age-group at each level of ROI. This revealed that the effect of age-group was significant for the time spent gazing to faces, $F(1, 81) = 7.84$, $p = .006$, $\eta_p^2 = .09$, but not bodies, $F(1, 81) = 2.21$, $p = .141$, $\eta_p^2 = .03$. As can be seen in Figure 2, older adults were gazing less at the faces of protagonists relative to younger adults, which indicates that older adults show biases in visual attention even when the stimuli are more naturalistic.

The two-way interaction between context and ROI was followed up by testing the simple effect of context at each level of ROI. This indicated that the effect of context was significant for gazing at faces, $F(1, 82) = 136.64$, $p < .001$, $\eta_p^2 = .63$, but not for gazing at bodies, $F(1, 82) = 1.68$, $p = .199$, $\eta_p^2 = .02$. As can be seen in Figure 3, participants were gazing less gazing at protagonists' faces in the enriched contextual condition compared to the minimal contextual condition.

Pearson correlation coefficients were calculated to assess whether visual biases, as indexed by percent fixation time towards the ROIs, correlated with ToM performance. As can be seen in Table 2, no significant associations emerged for either of the two age groups. This absence of any relationship between looking at faces or bodies and ToM ability suggests that the locus of visual attention does not play a major role in older adults' ability to decode mental states.

Study 2

Study 1 indicated that there are age differences in ToM capacity and the extent to which older adults attend to faces in naturalistic social scenes. However, the correlational data indicated that older and younger adults' gaze patterns were not related to ToM performance, which indicates that spending less time attending to faces does not appear to play a major role in older adults' difficulties on ToM tasks. Therefore, an important next step is to investigate other potential mechanisms that may underlie age-related difficulties in ToM capacity. ToM tasks require a series of processes that include basic processing and integrating of social information from the scenes, as well as more complex cognitive processes such as working memory and decision-making skills to choose between semantically complex labels. Therefore, one possible explanation for age differences in ToM capacity might involve difficulties in integrating and processing the most basic social cues about human agency and intentionality.

In order to accurately decode mental states, it is essential to detect and efficiently process basic cues about social agency (Frith & Frith, 2003), such as biological motion.

Indeed, previous studies have suggested that the ability to make inferences about other people's mental states develops from the ability to make inferences about others' actions (e.g., Blakemore & Decety, 2001; Castelli, Happé, Frith & Frith, 2000). Biological motion tasks are amongst the most widely used measures of basic social attention and typically require the processing of movement patterns from, for example, point-light displays of body parts which indicate that living agents are present (see Blake & Shiffrar, 2007 for a review). Indeed, even fleeting and partial information about biological motion can give strong clues to agency. Therefore, if older adults experience difficulties decoding basic social cues to agency, this could potentially impair more complex higher order social cognition skills such as ToM capacity.

Evidence to date indicates that older adults have some difficulties in decoding biological motion cues (e.g. Billino, Bremmer & Gegenfurtner, 2008; Insch, Bull, Phillips, Allen & Slessor, 2012; Legault, Troje & Faubert, 2012; Ruffman, Sullivan & Dittrich, 2009). Perhaps more interestingly, Phillips et al. (2011) found that age differences in biological motion processing partially mediated age effects in ToM processing. However, one limitation of Phillips et al.'s study is that they only included tasks that assessed one specific aspect of ToM – the ability to engage in false belief reasoning. In addition, the tasks included minimal social information. For instance, one task required the participant to read brief stories that described social situations and the other task required participants to watch soundless videos that depicted a person developing a false belief about the location of an object. Critically, these tasks did not require participants to integrate multiple social cues (i.e., body language, voice tone, facial expressions) nor did they include more subtle interactions between two or more people. Therefore, it's possible that efficiency of social attention (as indexed by biological motion perception accuracy) may contribute even more to older adults' ToM performance when the tasks are more socially complex and incorporate a greater number of social and emotional cues.

Another key variable which might explain age differences in ToM tasks is working memory capacity. Phillips et al. (2011) found that in addition to biological motion processing, working memory capacity also partially mediated age differences in false belief reasoning. This indicates that at least for very complex cognitive tasks involving social decisions, working memory capacity might be important. Indeed, consistent with the possibility that working memory may not have a universal role in ToM processing and instead may have a task-specific influence, another study showed that working memory did not explain age differences in a different aspect of ToM: detecting non-literal meaning in speech in relation to sarcasm (Phillips et al., 2015).

The goal of Study 2 was therefore to investigate whether age differences in ToM performance were mediated by variance in biological motion perception (as an index of social attention) and working memory capacity, using the same more naturalistic ToM assessments used in Study 1 (TASIT 2 & 3). If the efficiency of processing social information is critical for ToM capacity, then we would expect that biological motion perception will mediate age effects in ToM performance. However, if cognitive factors are more critical for ToM capacity, we would expect that working memory performance would mediate the age effects in ToM performance.

Method

Participants

In total 113 participants aged 19-86 took part in this study. Thirty-eight younger adults (M_{age} = 25.84, *SD* = 6.06, age range: 19-38 years), 40 middle-aged adults (M_{age} = 54.65, *SD* = 8.01, age range: 40-64 years), and 35 older adults (M_{age} = 74.09, *SD* = 4.62, age range: 65-86 years) participated in this study. Participants were recruited from local contacts and groups in Aberdeen, UK, and through the University of Aberdeen community, and were given a small payment for the testing session. All participants had English as their first language, 20/30 or corrected vision, and no history of substance abuse or presence of a psychiatric or neurological disorder. All older adults were screened for abnormal cognitive decline using the Mini Mental State Exam test (MMSE; Folstein, Folstein & McHugh, 1975). All scored above 26 (range 27-30, M = 28.83, SD = 1.04), which is the suggested cut-off for possible dementia in a well-educated sample (Mitchell, 2009).

As can be seen in Table 3, older adults had fewer years of education compared to middle-aged and older; while young adults had lower vocabulary scores than the other two groups. There were no differences between the groups with respect to self-rated health. The age groups also did not differ in gender, χ^2 (2) = 4.52, p = .105 (34.2%, 22.5% and 45.7% male, respectively).

Measures

Dynamic ToM tasks. As in Study 1, Parts 2 and 3 of the TASIT (Form A; McDonald et al., 2011) were used to assess ToM. Because Study 1 established that eye-movements were unrelated to ToM performance, and also because the explanatory variables of interest here were efficiency of social processing and working memory capacity, eye movements were not recorded in this study.

Efficiency of social attention: Biological motion. The action biological motion task reported by Phillips et al. (2011) was used to investigate ability to rapidly attend to social cues. Twelve point-light animations were presented to the participants in a random order, at a viewing distance of approximately 70cm, on a laptop computer. The actions were crawling, cycling, drinking, driving, jumping, playing pool, a tennis serve, rowing, saluting, sawing, sweeping and digging (Vanrie & Verfaillie, 2004). At the end of each animation participants said aloud which action they thought was being depicted in the trial. Because we were concerned about potential ceiling effects, participants were not provided with possible response options when making their response. The responses were coded as correct or incorrect and the dependent variable was the percentage of actions correctly reported.

Working memory measure. The N-back task from Phillips et al. (2011) was used investigate ability to update and monitor working memory. This task involved the presentation of a sequence of numbers (ranging from 1 to 9) individually in the center of the screen. During a 1-back practice block of trials, participants had to monitor stimuli and decide whether the presented number was the same as the number that had been presented immediately before by pressing one of two keys labelled ‘yes’ and ‘no’ on a response box. The 2-back updating condition consisted of 38 trials where participants had to report whether the number presented was the same as the number before the last (i.e., presented two trials before it) by pressing either the yes or no key. Each stimulus remained on screen until the participant had made a response. This was then followed by an inter-stimulus interval of 800ms. Percentage accuracy on the 2-back condition was used as the dependent measure. Note that ten participants did not complete this task leading to an n of 103.

Procedure

The study received approval from The University of Aberdeen Psychology Ethics Committee. All participants were tested individually across two separate testing sessions, with a gap of roughly one month between sessions. At the first session participants were given an information sheet, followed by a demographic questionnaire. The MMSE was then administered to older adults. Next, all participants completed the Mill Hill, biological motion and working memory task as well as some other tasks not considered here. In the second session they completed the TASIT tasks as well as questionnaires not further analysed here. Participants were then provided with reimbursement and debriefed.

Analyses

Given the continuous nature of age range in this sample, initial correlations were calculated between age and assessments of ToM ability, biological motion processing and working memory capacity. To investigate the potential mediating role of social attention and working memory in age differences in ToM ability, initially correlations between the key measures were explored, followed by regression analyses specified by PROCESS (Hayes, 2013) to test for mediation effects.

Results and Discussion

Correlations Between Age, ToM Performance, Social Attention and Working Memory.

Table 4 reports correlations between age and the measures of ToM, social attention (as assessed by biological motion) and working memory. In addition, the relationships between the ToM tasks, biological motion and working memory are reported. Consistent with the results from Study 1, age negatively correlated with scores on both TASIT ToM tasks, however this correlation was strongest in TASIT 3 (enriched context). A *z*-test (Lee & Preacher, 2013) was used to compare the strength of correlations, showing that age was more strongly associated with poorer performance on the enriched context compared to the minimal context task, $z = -2.24$, $p < .05$. Age was also negatively correlated with social attention and working memory performance. Next, the associations between these potential mediator variables and the ToM tasks was explored. ToM performance was reliably related to social attention as assessed by biological motion detection, while only performance on the TASIT 2 (minimal context) ToM task was related to working memory performance.

Mediation Analyses

To investigate whether age differences in ToM capacity might be mediated by measures of social attention (biological motion) or working memory (N-back), mediation analyses were carried out. The methods outlined by Hayes (2013) were followed using bootstrap procedures with 5000 resamples to test the significance of any mediation effects.

The first mediation analysis explored whether social attention mediated age differences in ToM capacity as assessed by the TASIT 2 (minimal context) task (see Figure 4i). Biological motion detection was a significant predictor of ToM performance when entered alongside age, and was a significant mediator of the age-ToM relationship (bias corrected confidence intervals from -.1388 to -.0207 do not include zero). Here the total R^2 for the regression model was .100, and the mediation effect of biological motion explained nearly all of the age differences in ToM capacity. The direct age effect on TASIT 2 performance is no longer significant once biological motion is included in the model. The second mediation analysis looked at predicting TASIT 3 (enriched context) performance (see Figure 4ii) from age and biological motion performance. Biological motion was a significant predictor of ToM performance when entered alongside age in the regression analysis. It was also a significant mediator of age differences in ToM (bias corrected confidence intervals from -.1555 to -.0342 do not include zero). The total R^2 for the regression model was .230, and the mediation effect of biological motion explained nearly all of the age variation. Again, the effect of age on TASIT 3 was no longer significant once biological motion was included in the model. In summary, age differences in ToM were fully explained in regression models by variability in biological motion perception.

A further set of regression analyses explored the potential mediating role of working memory in age differences in ToM performance. The first looked at whether age differences in ToM performance on the TASIT 2 (minimal context) task were mediated by working memory capacity (see Figure 5i). Working memory was not a significant predictor of ToM performance when entered alongside age, and did not significantly mediate the age-ToM relationship (bias corrected confidence intervals from -.0653 to .0033 include zero). Here the total R^2 for the regression model was .111. A final mediation analysis looked at the same model predicting TASIT 3 (enriched context) performance (see Figure 5ii). Working memory did not predict ToM performance when entered alongside age in the regression analysis, and did not mediate age differences in ToM (bias corrected confidence intervals from -.0322 to .0348 include zero). The total R2 for the regression model was .168. No significant mediation effects were found, indicating that age differences in ToM could not even partially be explained by common variance with working memory performance indexed by the n-back task.

General Discussion

The two experiments reported here provide novel insights into ToM processing in older age and the potential mechanisms that may explain why older adults show difficulties on ToM tasks. One of the primary aims of both studies was to characterise the pattern and magnitude of age differences in ToM understanding in response to more naturalistic social scenes that represent everyday social encounters. The results from both studies showed that older adults had greater difficulties relative to their younger counterparts decoding mental states and intentions with the provision of more naturalistic ToM tasks. These findings suggest that age-related difficulties observed on traditional lab-based tasks may extend to more complex and realistic social situations. These findings are not in line with prior studies that have shown that older adults perform better on tasks that are more meaningful and relevant (Germain & Hess, 2007; Hess et al., 2001; 2005; 2009).

Another important question of interest was whether older adults would benefit from increased contextual information presented during the social scene stimuli. Prior studies in both the cognitive ageing and emotion perception literatures have shown that older adults particularly benefit from additional contextual cues (Lindenberger & Mayr, 2014; Noh & Isaacowitz, 2015; Noh & Isaacowitz, 2013; Smith et al., 1998), but this has not previously been applied to ToM tasks. In both of our studies, ToM tasks were used that differed in the amount of contextual information available. Whereas TASIT 2 included minimal contextual information, TASIT 3 included more enriched context by providing additional verbal and visual cues to help facilitate mental state understanding. If older adults benefit from context, this should result in reduced age differences on the enriched compared to minimal context tasks. In fact, the results showed the reverse, with age-related difficulties greater on the enriched context task. Importantly, both the Australian (Study 1) and UK (Study 2) samples showed similar patterns of age effects on ToM performance, with larger effects observed with the provision of increased contextual information.

The most parsimonious explanation for these findings, robust across two separate studies, is that the added complexity of the enriched context condition made it more difficult for older adults to effectively use the contextual cues available. Indeed, while Noh and Isaacowitz (2013) showed that contextual cues are beneficial for older adults' emotion recognition abilities, they used simple stimuli (static photographs of e.g. an angry facial expression), only incorporated one contextual cue (e.g., a fist to indicate anger), and required judgments of basic emotions. This contrasts with the demands of TASIT 3, which requires participants to track the mental states of two or more protagonists, while also attending to visual and verbal contextual cues embedded in dynamic background stimuli that reveal the true state of affairs. Because this latter type of task not only imposes greater demands on higher level mental state reasoning processes, but also because it is more perceptually complex, older adults may have found it difficult to process the multiple paralinguistic cues in the scenes and consequently may have been less able to use the contextual cues available.

Overall, these findings indicate that there may be a limit to how much context is beneficial for older adults. If the stimuli are simple with basic contextual information, context may be helpful, but if the stimuli are complex with multiple contextual and paralinguistic cues, context may not provide any added benefit, and may even prove detrimental. Further research is now needed to gain a clearer and more nuanced understanding of how stimulus complexity impacts on social cognitive ability, and whether the age differences identified in response to the more naturalistic ToM scenarios map on to real life social interactions, where contexts and interactions may well be complex.

Possible factors which might explain older adults' ToM difficulties

In addition to providing novel insights into the nature and magnitude of age differences in ToM capacity, both Study 1 and Study 2 included additional assessments to address potential mechanisms that may explain why older adults experience difficulties on ToM tasks. Study 1 measured participants' eye-gaze patterns while they completed the ToM tasks to assess whether older adults show visual biases when viewing naturalistic social scene stimuli. As noted previously, eye movements were important to assess because prior studies have shown that older adults look less at the most emotionally rich areas of a face (i.e., the eye region, Circelli et al., 2013; Grainger et al., 2017; Sullivan et al., 2007), but have identified these effects using stimuli with limited ecological validity (predominantly static images).

Adding to these prior studies and consistent with the only other study to measure visual biases towards dynamic social interactions (Vicaria et al., 2015), Study 1 revealed that older relative to younger adults focused their attention less towards the faces of the protagonists. However, no differences were observed in looking at the bodies of the protagonists. Importantly, this age effect did not differ for the minimal and enriched contextual conditions. Thus, this finding, taken along with the eye-tracking studies cited above, indicates that older adults display a bias in visual processing of social stimuli that is not dependent on the ecological validity or complexity of the social cues. It has previously been argued that older adults' reduced attention towards faces may reflect a difficulty filtering out non-social background components of the scene and a greater tendency to spontaneously encode irrelevant information (Ngo & Isaacowitz, 2015). By identifying age-related biases away from the most socially rich components of the stimuli and thus towards non-social information in both ToM tasks, the data from this study support this interpretation.

However, perhaps most critically, no associations were identified between visual biases and ToM accuracy scores for either age-group, suggesting that although older adults have a tendency to look less at the most socially rich information, these biases are not impacting on their ability to engage in mental state understanding. This failure to find a relationship between visual biases and ToM ability is not consistent with Vicaria et al., (2015) who found that gazing to faces impacted rapport judging ability differently for young and older adults. However, these findings are consistent with most previous studies that have found no relationship between eye-gaze patterns and emotion recognition ability (e.g., Circelli et al., 2013; Grainger et al., 2017; He et al., 2011; Sullivan et al., 2007), and suggests that although age differences are evident in both attentional focus and ToM ability, there are likely different mechanisms that contribute to these effects.

Because ToM ability was not associated with visual biases in Study 1, Study 2 examined whether efficiency of social attention (as indexed by biological motion perception ability) might explain age differences in ToM. As noted earlier, this methodological approach was used because in order to decode emotions and other mental states, it is essential to detect and process basic cues about social agency (Frith & Frith, 2003), such as biological motion. Supporting our predictions, the results showed that social attention explained almost all of the age-related variance in both ToM tasks (i.e., TASIT 2 & 3), which suggests that the ability to detect and rapidly process low level cues to social agency may be important in age differences in higher order social cognitive processes, such as ToM. We propose that older adults' less efficient processing of these basic cues providing social information may incrementally result in impaired ability to integrate more complex social cues from dynamic multimodal scenes. These data are also consistent with Phillips et al.'s (2011) findings, in which it was shown that social attention (again as indexed by biological motion perception) partially mediated the age-related variance in much less contextualized false belief understanding tasks. As discussed earlier, the tasks used in Phillips et al. were single modality (i.e., reading story passages and watching silent videos) and therefore required less rich social processing compared to the TASIT stimuli used in our study.

It is likely that social attention accounted for more age-related variance in ToM processing in our study because the social processing demands of the TASIT tasks were greater. That is, participants had to attend to and integrate a greater number of social cues (i.e., voice tone, body language, facial expressions) in order to perform well on the task. Together, these findings indicate that the efficiency of social attention plays an important role in age differences in higher order social cognitive processes such as ToM capacity, however this role appears to be much greater when the task or situation includes more complex stimuli that incorporate multiple social cues.

In sum, the results from Study 2 indicate that older adults have difficulties detecting, integrating and processing social cues (low-level perceptual information) about human actions, which is potentially interfering with higher-order mental state processing. Thus, efficiency in integrating and processing multiple social cues, rather than the locus of visual attention, could be a critical skill that may underlie the ability to engage in mental state understanding. This likely reflects the situation in everyday social interactions, which require the continuous processing of subtle social signals that rapidly appear and disappear.

In addition to examining whether more basic social attention could explain age differences in ToM, Study 2 also assessed working memory capacity to examine whether cognitive capacity more generally could explain the age differences in ToM performance. In contrast to social attention, working memory did not explain any of the age-related variance in ToM performance. Phillips et al. (2011) found that updating working memory partially mediated the age effects in false belief reasoning, however several other prior studies have reported that age-related difficulties in other aspects of ToM performance continue to remain after controlling for working memory (Bernstein, Thornton, & Sommerville, 2011; Cavallini, Lecce, Bottiroli, Palladino, & Pagnin, 2013; German & Hehman, 2006; Phillips et al., 2015). It seems likely that the characteristics of the ToM task may be an important determinant of which specific cognitive resources are involved. For instance, the ability to judge and hold two or more different mental states simultaneously in one's mind throughout a series of events (as required in false belief reasoning tasks, Phillips et al., 2011) may be more dependent on working memory capacity compared to detecting when someone is being sincere or sarcastic (as required in the TASIT task used in this study). Nevertheless, the weight of evidence to date suggests that working memory capacity is not a major contributor to the age effects in most aspects of ToM, and instead the evidence points to efficiency of social attention being an important contributor to age effects in ToM.

Limitations and directions for future research

While the two studies reported here have provided novel insights into factors which might explain age effects in ToM processing, there are a number of questions that need to be explored in future research. Firstly, although the TASIT was used because it includes naturalistic social scene stimuli, these stimuli are limited by the fact that they are not genuine social interactions and do not include familiar social targets. Indeed, it has recently been shown that age differences in emotion recognition accuracy are attenuated when viewing dynamic facial expressions of a familiar romantic partner versus a same-age stranger (i.e., a familiarity manipulation; Stanley & Isaacowitz, 2015). Since the TASIT stimuli are enacted by unfamiliar actors, it is possible that the age-related ToM difficulties identified on the TASIT may overestimate the difficulties that older adults experience in their everyday lives with family and friends. In addition, the TASIT stimuli primarily include young and middle-aged actors and therefore the scenarios may potentially be more relevant for younger relative to older adults. The age of protagonists has not generally been addressed in ToM stimuli, though it is worth noting that there is little evidence of an own-age bias in older adults' performance in other social domains such as emotion perception (Ebner & Johnson, 2009) or gaze following (Slessor, Laird, Phillips, Bull, & Filippou, 2010). Nevertheless, it is still an important next step in this literature to develop age-appropriate naturalistic tasks of ToM that are suitable for both older and younger adults.

In regards to the eye-tracking component in Study 1, one key limitation was that we did not have pre-defined ROI's to examine what aspects of the background participants were attending to. This is because the background cues were often not consistent in size across the trials and because half of the trials in TASIT 3 included verbal contextual information (i.e., in the form of an epilogue or prologue), which is difficult to capture visually. Thus, while we know that older adults were looking at the background scene information to a greater extent than young in both components of the TASIT, it is unclear what particular aspects of the background information they were focusing on. Future research could more directly manipulate the way in which background contextual information was provided in order to tease apart age differences in prioritising particular types of contextual cue. Also, the addition of contextual information could be titrated to discover whether there are age differences in the optimal amount of context.

Finally, although we did not find a relationship between eye-gaze patterns and ToM performance, future research is needed to further understand why older adults consistently show visual biases during social cognitive processing. One possibility is that visual biases in older age may reflect an adaptive process to compensate for reductions in other sensory functions (e.g., hearing). Further studies are now needed to investigate other factors (e.g., task type) that may be contributing to visual biases in older age. Moreover, it will also be important in future studies to clarify whether these age-related visual biases have any direct or indirect consequences on social cognitive abilities.

Conclusion

These data provide an important step in understanding the extent to which older adults experience difficulties in ToM capacity. Across two separate studies with independent samples, age-related difficulties were observed on more naturalistic ToM measures, indicating that previously reported age-related difficulties in ToM tasks with limited ecological validity also extend to tasks that are more closely representative of everyday social interactions. In addition, in both studies, older adults performed more poorly on the ToM task that included more enriched context, supporting the hypothesis that too much information may overload the processing capacity of older people when decoding mental states.

The current work also provides novel insights into the potential mechanisms that may explain age-related difficulties in ToM capacity. Although older adults looked less than young at socially rich areas of the scenes (i.e., faces), these biases were not related to their overall ToM accuracy. This indicates that age-related biases in the allocation of attention seem unlikely to explain age differences in decoding intentions. In contrast, the ability to integrate and decode basic biological motions cues did explain the age variance in ToM performance, which suggests that it may not be the locus of visual attention that is critical to ToM ability, but instead the efficiency of processing social cues. Interestingly, working memory capacity did not explain any age-related variance in ToM performance, suggesting age differences in social cognition are likely to be somewhat independent of more general age-related declines in cognition.

Acknowledgments

This work was supported by a Discovery Project grant (DP150100302) from the Australian Research Council awarded to Henry & Phillips.

References

Achim, A. M., Guitton, M., Jackson, P. L., Boutin, A., & Monetta, L. (2013). On what ground do we mentalize? Charateristics of current tasks and sources of information that contribute to mentalizing judgments. *Psychological Assessment, 25,* 117-126. doi: 10.1037/a0029137.

Bernstein, D. M., Thornton, W. L., & Sommerville, J. A. (2001). Theory of mind through the ages: older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Experimental Aging Research, 37,* 481-502. doi: 10.1080/0361073X.2011.619466

Billino, J., Bremmer, F., & Gegenfurtner, K. R. (2008). Differential ageing of motion processing mechanisms: Evidence against general perceptual decline. *Vision Research, 48,* 1254-1261. doi: 10.1016/j.visres.2008.02.014

Blake, R., & Shiffrar, M. (2007). Perception of human motion. *Annual Review of Psychology, 58,* 47-73. doi: 10.1146annurev.psych.57.102904.190152

Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience, 2,* 561-567. doi: 10.1038/35086023

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62,* 193-217. doi: 10.1037/h0047470.

Burdon, P., Dipper, L., & Cocks, N. (2016). Exploration of older and younger British adults' performance on The Awareness of Social Inference Test (TASIT). *International Journal of Language and Communication Disorders, 51,* 589-593. doi: 10.1111/1460-6984.12233

Carstensen, L. L., Mikels, J. A., & Mather, M. (2006). Ageing and the intersection of cognition, motivation and emotion. *Handbook of the Psychology of Aging, 6,* 343-362.

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: a functional imageing study of perception and interpretation of complex intentional movement patterns. *Neuroimage, 12,* 314-325. doi: 10.1006/nimg.2000.0612

Cavallini, E., Lecce, S., Bottiroli, S., Palladino, P., & Pagnin, A. (2013). Beyond false belief: theory of mind in young, young-old, and old-old adults. *International Journal of Aging and Human Development, 76*, 181-198. doi:10.2190/AG.76.3.a

Circelli, K. S., Clark, U. S., & Cronin-Golomb, A. (2013). Visual scanning patterns and executive function in relation to facial emotion recognition in ageing. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 20,* 148-173. doi: 10.1080/13825585.2012.675427

Craik, F. I. M, & Byrd, M. (1982). Aging and cognitive deficits: the role of attentional resources. In F. I. M. Craik & S. Trehub (Eds.). *Aging and Cognitive Processes* (p.191). New York: Plenum.

Ebner, N. C., & Johnson, M. K. (2009). Young and older emotional faces: are there age group differences in expression identification and memory? *Emotion, 9*, 329-339. doi: 10.1037/a0015179.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-Mental State": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198. http://dx .doi.org/10.1016/0022-3956(75)90026-6

Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, Series B, 358*, 459-473. doi: 10.1098/rstb.2002.1218

Germain, C. M., & Hess, T. M. (2007). Motivational influences on controlled processing: Moderating distractibility in older adults. *Aging, Neuropsychology, and Cognition, 14*, 462-486. doi: 10.1080/13825580600611302

German, T. P., & Hehman, J. A. (2005). Representational and executive selection resources in ‘theory of mind’: Evidence from compromised belief-desire reasoning in old age. *Cognition, 101,* 129-152. doi: 10.1016/j.cognition.2005.05.007

Glisky, E. L. (2007). Changes in cognitive function in human ageing. In Riddle, D. R. (Ed). *Brain Aging: Models, Methods, and Mechanisms* (Chapter 1). Boca Raton: CRC Press/Taylor & Francis

Grainger, S., Henry, J., Phillips, L., Vanman, E, & Allen, R. (2017). Age deficits in facial affect recognition: The influence of dynamic cues. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 72,* 622-632. doi: 10.1093/geronb/gbv100

Grühn, D., Rebucal, K., Diehl, M., Lumley, M., & Labouvie-Vief, G. (2008). Empathy across the lifespan: Longitudinal and experience-sampling findings. *Emotion, 8,* 753-765. doi: 10.1037/a0014123

Gutchess, A. H., Hebrank, A., Sutton, B. P., Leshikar, E., Chee, M. W. L., Tan, J. C., . . . Park, D. C. (2007). Contextual interference in recognition memory with age. *NeuroImage, 35*, 1338-1347. doi: 10.1016/j.neuroimage.2007.01.043

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York: Guilford Press.

He, Y., Ebner, N. C., & Johnson, M. K. (2011). What predicts the own-age bias in face recognition memory. *Social Cognition, 29,* 97-109. doi: 10.1521/soco.2011.29.1.97

Henry, J. D., Phillips, L. H., Ruffman, T., & Bailey, P. E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28,* 826-839. doi: 10.1037/a0030677

Henry, J. D., von Hippel, W., Molenberghs, P., Lee, T., & Sachdev, P. S. (2016). Clinical assessment of social cognitive function in neurological disorders. *Nature Reviews. Neurology, 12,* 28. doi:10.1038/nrneurol.2015.229

Hess, T. M. (2006). Adaptive aspects of social cognitive functioning in adulthood: Age-related goal and knowledge influences. *Social Cognition*, *24*, 279-309.

Hess, T. M., Germain, C. M., Rosenberg, D. C., Leclerc, C. M., & Hodges, E. A. (2005). Ageing-related selectivity and susceptibility to irrelevant affective information in the construction of attitudes. *Aging, Neuropsychology, and Cognition*, *12*, 149-174. doi: 10.1080/13825580590925170

Hess, T. M., Germain, C. M., Swaim, E. L., & Osowski, N. L. (2009). Ageing and selective engagement: The moderating impact of motivation on older adults' resource utilization. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *64*, 447-456. doi: 10.1093/geronb/gbp020

Hess, T. M., Rosenberg, D. C., & Waters, S. J. (2001). Motivation and representational processes in adulthood: the effects of social accountability and information relevance. *Psychology and Aging*, *16*, 629-642. doi: 10.1037/0882-7974.16.4.629

Insch, P. M., Bull, R., Phillips, L. H., Allen, R., & Slessor, G. (2012). Adult ageing, processing style, and the perception of biological motion. *Experimental Aging Research, 38*, 169-185. doi: 10.1080/0361073X.2012.660030

Isaacowitz, D. M., & Stanley, J. T. (2011). Bringing an ecological perspective to the study of ageing and recognition of emotional facial expressions: Past, current, and future methods. *Journal of Nonverbal Behavior, 35,* 261-278. doi: 10.1007/s10919-011-0113-6

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006). Is there an age-related positivity effect in visual attention? A comparison of two methodologies. *Emotion, 6*, 511-516. doi:10.1037/1528-3542.6.3.511

Isaacowitz, D. M., Wadlinger, H. A., Goren, D., & Wilson, H. R. (2006). Selective preference in visual fixation away from negative images in old age? An eye-tracking study. *Psychology and Aging, 21,*40-48. doi: 10.1037/0882-7974.21.1.40

Kunzmann, U, & Isaacowitz, D. (2017). Emotional Ageing: Taking the immediate context seriously. *Research in Human Development, 14,* 182-199. doi: 10.1080/15427609.2017.1340048

Labouvie-Vief, G., Diehl, M., Jain, E., & Zhang, F. (2007). Six year change in affect optimization and affect complexity across the adult lifespan: A further investigation. *Psychology and Aging, 22,* 738-751. doi: 10.1037/0882-7974.22.4.738

Lee, I. A., & Preacher, K. J. (2013, September). Calculation for the test of the difference between two dependent correlations with one variable in common [Computer software]. Available from http://quantpsy.org.

Legault, I., Troje, N. F., & Faubert, J. (2012). Healthy older observers cannot use biological-motion point-light information efficiently within 4 m of themselves. *Iperception, 3,* 104-111. doi: 10.1068/i0485

Lindenberger, U., & Mayr, U. (2014). Cognitive aging: is there a dark side to environmental support? *Trends in Cognitive Sciences, 18,* 7-15. doi: 10.1016/j.tics.2013.10.006

Manor, B. R., & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of Neuroscience Methods, 128,* 85-93. doi: 10.1016/S0165-0270(03)00151-1

McDonald, S., Flanagan, S., & Rollins, J. (2011). The awareness of social inference test - revised (TASIT). Sydney: Pearson Education Ltd.

Mioshi, E., Dawson, K., Mitchell, J., Arnold, R., & Hodges, J. R. (2006). The Addenbrooke's cognitive examination revised (ACE-R): A brief cognitive test battery for dementia screening. *International Journal of Geriatric Psychiatry, 21*, 1078-1085. doi: 10.1002/gps.1610

Mitchell, A. J. (2009). A meta-analysis of the accuracy of the mini-mental state examination in the detection of dementia and mild cognitive impairment. *Journal of Psychiatric Research, 43,* 411–431. doi:10.1016/j.jpsychires.2008.04.014

Moran, J. M. (2013). Lifespan development: The effects of typical ageing on theory of mind. *Behavioural Brain Research, 237,* 32-40. doi: 10.1016/j.bbr.2012.09.020

Murphy, N. A., & Isaacowitz, D. M. (2010). Age effects and gaze patterns in recognizing emotional expressions: An in-depth look at gaze measures and covariates. *Cognition & Emotion, 24,* 436-452. doi: 10.1080/02699930802664623

Nelson, H. E. (1982). *National Adult Reading Test*. Windsor, UK: NFER-Nelson.

Ngo, N., & Isaacowitz, D. M. (2015). Use of context in emotion perception: The role of top-down control, cue type, and perceiver's age. *Emotion, 15*, 292-302. doi: 10.1037/emo0000062

Noh, S. R., & Isaacowitz, D. M. (2013). Emotional faces in context: Age differences in recognition accuracy and scanning patterns. *Emotion, 13*, 238-249. doi: 10.1037/a0030234

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: from measures to mechanisms. *Trends in Cognitive Sciences, 15*, 152-159. doi: 10.1016/j.tics.2011.01.005

Phillips, L. H., Allen, R., Bull, A., Kliegal, M., Channon, S. (2015). Older adults have difficulty decoding sarcasm. *Developmental Psychology, 51,* 1840-1852. doi: 10.1037/dev0000063.

Phillips, L. H., Bull, R., Allen, R., Insch, P., Burr, K., & Ogg, W. (2011). Lifespan ageing and belief reasoning: Influences of executive function and social cue decoding. *Cognition, 120*, 236-247. doi:10.1016/j.cognition.2011.05.003

Phillips, L. H., Scott, C., Henry, J. D., Mowat, D., & Bell, J. S. (2010). Emotion perception in Alzheimer's disease and mood disorders in old age. *Psychology and Aging, 25*, 38-47. doi: 10.1037/a0017369

Raz, N., Lindenberger, U., Rodrigue, K. M., Kennedy, K. M., Head, D., Williamson, A., . . . Acker, J. D. (2005). Regional brain changes in aging healthy adults: General trends, individual differences and modifiers. *Cerebral Cortex, 15*, 1676-1689. doi: 10.1093/cercor/bhi044

Riby, D., & Hancock, P. J. B. (2009). Looking at movies and cartoons: eye-tracking evidence from Williams syndrome and autism. *Journal of Intellectual Disability Research, 53*, 169-181. doi: 10.1111/j.1365-2788.2008.01142.x

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and ageing: Implications for neuropsychological models of ageing. *Neuroscience and Biobehavioral Reviews, 32*, 863-881. doi: 10.1016/j.neubiorev.2008.01.001

Ruffman, T., Sullivan, S., & Dittrich, W. (2009). Older adults' recognition of bodily and auditory expressions of emotion. *Psychology and Aging, 24*, 614-622. doi: 10.1037/a0016356

Salthouse, T. A. (1995). Processing capacity and its role on the relations between age and memory. In F. E.Weinert & W. Schneider (Eds.). *Memory Performance and Competencies: Issues in Growth and Development* (p. 111). Hillsdale, NJ: Erlbaum

SensoMotoric Instruments GmbH. (2015). *BeGaze™* (Version 3.5) [Software].

SensoMotoric Instruments GmbH. (2015). *Experiment Centre™* (Version 3.5) [Software].

SensoMotoric Instruments GmbH. (2011). *iView X* (Version 2.8) [Apparatus and software].

Slessor, G., Laird, G., Phillips, L. H., Bull, R., & Filippou, D. (2010). Age-related differences in gaze following: Does the age of the face matter? *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *65,* 536-541. doi:10.1093/geronb/gbq038

Smith, A. D., Park, D. C., Earles, J. L. K., Shaw, R. J., & Whiting, W. L. (1998). Age differences in context integration in memory. *Psychology and Aging, 13,* 21-28. doi: 10.1037/0882-7974.13.1.21

Spieler, D. H., Mayr, U., & Lagrone, S. (2006). Outsourcing cognitive control to the environment: Adult age differences in the use of task cues. *Psychonomic Bulletin & Review, 13*, 787-793. doi: 10.3758/BF03193998

Stanley, J. T., & Isaacowitz, D. M. (2015). Caring more and knowing more reduces age-related differences in emotion perception. *Psychology and Aging, 30,* 383-395. doi:10.1037/pag0000028

Sullivan, S., Campbell, A., Hutton, S. B., & Ruffman, T. (2015). What's good for the goose is not good for the gander: Age and gender differences in scanning emotion faces. *Journal of Gerontology: Psychological Sciences,* advance online publication. doi: 10.1093/geronb/gbv033.

Sullivan, S., Ruffman, T., & Hutton, S. B. (2007). Age differences in emotion recognition skills and the visual scanning of emotion faces. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences, 62,* 53-60. doi: 10.1093/geronb/62.1.P53

Vanrie, J., & Verfaillie, K. (2001). Perception of biological motion: a stimulus set of human point-light actions. *Behavior Research Methods, Instruments & Computers,* 625-629. doi: 10.3758/BF03206542

Vicaria, I. M., Bernieri, F. J., & Isaacowtiz, D. M. (2015). Perceptions of rapport across the life span: Gaze patterns and judgment accuracy. *Psychology and Aging, 30,* 396-406. doi: 10.1037/pag0000019

Wong, B., Cronin-Golomb, A., & Neargarder, S. (2005). Patterns of visual scanning as predictors of emotion identification in normal ageing. *Neuropsychology*, *19*, 739-749. doi: 10.1037/0894-4105.19.6.739

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67,* 361-370. doi: 10.1111/j.1600-0447.1983.tb09716.x

Table 1
Participant Characteristics

Measure	Younger		Older		Inferential Statistics			
	M	*SD*	*M*	*SD*	*t*	*df*	*p*	*d*
Education	13.72	1.77	13.42	3.27	0.57	93	.570	.11
Self-rated health	4.10	0.70	4.30	0.76	1.17	95	.244	.27
NART	110.31	3.88	116.27	5.21	6.23	91	<.001	1.30
HADS total	2.08	.40	2.44	0.29	5.00	92	<.001	1.03

Note. Self-rated health was indexed by responses on a 5-point scale, with 1 indicating *very poor* and 5 indicating *very good.* NART refers to the *National Adult Reading Test*. HADS refers to the Hospital Anxiety and Depression Scale, and HADS total provides an overall index of negative affect. Note that a higher score for HADS total reflects greater positive affect. Effect sizes of group differences are expressed as Cohen's *d.*

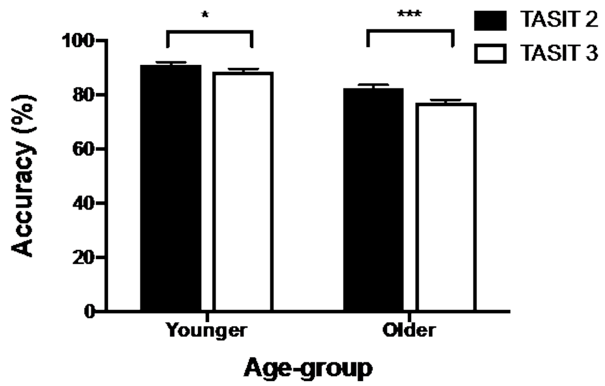

Figure 1. Accuracy scores (%) for younger and older adults on TASIT 2 (minimal context) and TASIT 3 (enriched context). Error bars represent standard error of the mean.
* = p < .05, *** = p < .001

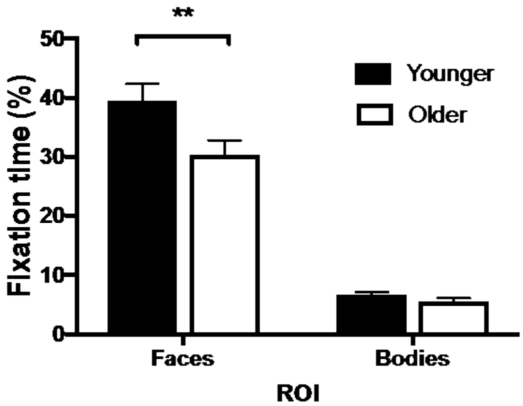

Figure 2. Mean fixation time (%) for older and younger adults towards each ROI, collapsed across task-type (TASIT 2, TASIT 3) $**p < .01$

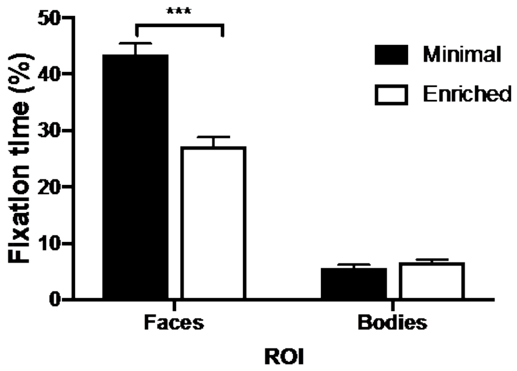

Figure 3. Mean fixation time (%) towards each ROI in the minimal and enriched ToM tasks, collapsed across age-group ***p < .001

Table 2.

Within-task Pearson correlations between fixation time to the ROI's (faces, bodies) and TASIT 2 & 3 accuracy separately for younger and older adults.

	ROI: Faces			ROI: Bodies		
	n	*r*	*p*	*n*	*r*	*p*
Younger						
Correlation with accuracy: TASIT 2	47	.043	.775	47	.196	.186
Correlation with accuracy: TASIT 3	46	-.031	.838	46	.104	.490
Older						
Correlation with accuracy: TASIT 2	38	.150	.368	38	.013	.938
Correlation with accuracy: TASIT 3	39	.258	.113	39	-.044	.789

Table 3.

Participant Characteristics

Measure	Younger		Middle-aged		Older		Inferential Statistics		
	M	*SD*	*M*	*SD*	*M*	*SD*	*F(112)*	*p*	*f*
Education (years)	16.3	2.34	16.8	4.40	13.3	4.22	8.84	<.001	0.76
Self-rated health	7.47	1.08	7.20	1.18	7.09	1.22	1.09	.340	0.14
Mill Hill	18.29	4.39	23.40	4.05	22.86	4.43	16.37	<.001	1.11

Note. Self-rated health was indexed by responses on a 9-point Likert scale, with 1 indicating *very poor* and 9 indicating *very good.* Mill Hill is the *Mill Hill Vocabulary Test*. Effect sizes of group differences are expressed as Cohen's *f*.

Table 4.

Pearson correlations between age, ToM performance, social attention and working memory, and descriptive statistics for the tasks. Note that all descriptive statistics are reported as percentage accuracy.

	TASIT 2 ToM-minimal	TASIT 3 ToM-enriched	Biological motion accuracy	Working memory n-back accuracy
Age	-.211 *	-.380 ***	-.569 ***	-.306 **
Biological motion accuracy	.314 **	.456 ***	-	.180
Working memory n-back accuracy	.223 *	.125	.180	-
M	85.4	79.9	62.6	92.9
SD	10.28	9.31	18.97	8.04

* $p < .05$, ** $p < .01$, *** $p < .001$.

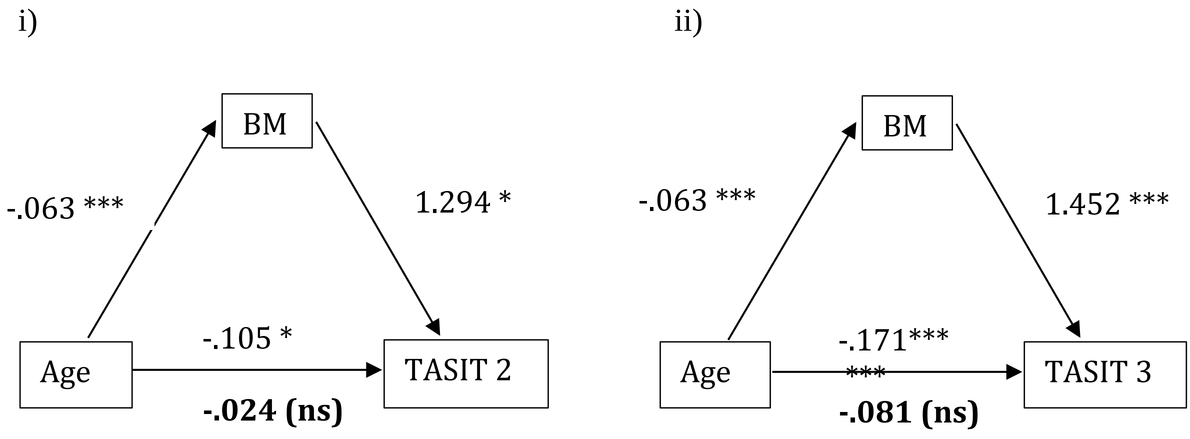

Figure 4. Path diagrams of mediation analyses examining biological motion detection (BM) as a possible mediator of age differences in ToM as assessed by the (i) TASIT 2 and (ii) TASIT 3. The numbers are uncorrected path coefficients as recommended in Preacher and Hayes (2008). The coefficients above the pathway between age and the social cognition measures are the total effects, and the bold figure in each case below the pathway is the direct age effect on social cognition once biological motion is included in the model. The significance of the mediation effect is determined by a bootstrapping procedure, see text for results. Note n = 113. BM = biological motion task. * p < .05, ** p < .01, *** p < .001.

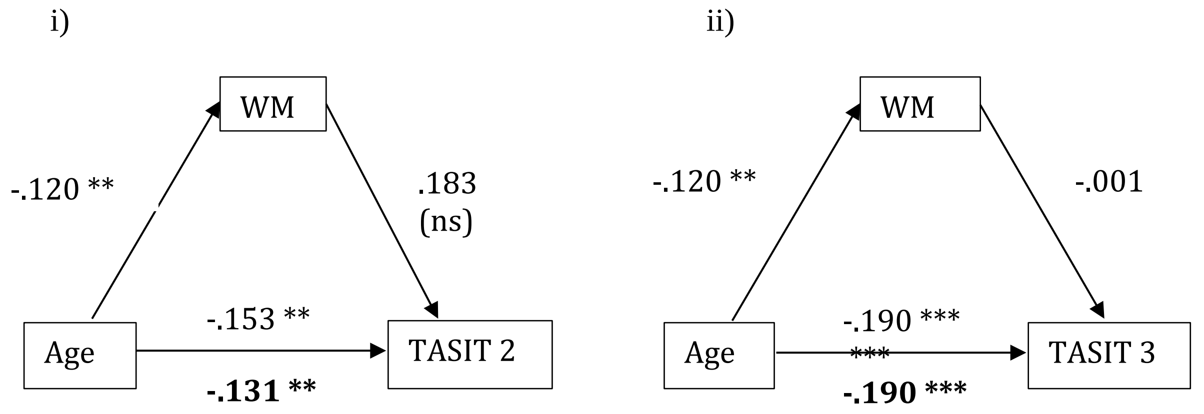

Figure 5. Path diagrams of mediation analyses examining working memory capacity (WM) as a possible mediator of age differences in ToM as assessed by the (i) TASIT 2 and (ii) TASIT 3. The numbers are uncorrected path coefficients as recommended in Preacher and Hayes (2008). The coefficients above the pathway between age and the social cognition measures are the total effects, and the bold figure in each case below the pathway is the direct age effect on social cognition once biological motion is included in the model. The significance of the mediation effect is determined by a bootstrapping procedure, see text for results. Note n = 103. WM = working memory capacity. * $p < .05$, ** $p < .01$, *** $p < .001$.